SALIENCY BASED AUTOMATED

IMAGE CROPPING

By

MOUNIKA RAMINENI

Bachelor of Science in Information Technology

MVGR College of Engineering

Vizianagaram, India

2015

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 2017

SALIENCY BASED AUTOMATED

IMAGE CROPPING

Thesis Approved:

Christopher John Crick

Thesis Adviser

K. M. George

Eric Chan-Tin

Name: MOUNIKA RAMINENI

Date of Degree: DECEMBER, 2017

Title of Study: SALIENCY BASED AUTOMATIC IMAGE CROPPING

Major Field: COMPUTER SCIENCE

Abstract: Automatic Image Cropping have been enormously popular in the fields of photography, printing industries and in many other fields. It can be usually done in two ways; attention based and aesthetic based. Attention based approach mainly focuses on the important subject in the image whereas aesthetic based approach focuses on the attractiveness of the image. In this project, we proposed a novel attention based approach, which crops out the distracting objects from the image. In our approach, spatially weighted dissimilarity saliency model is used for object detection. In contrast to many methods, cropping of the detected object is done based on the centroid which is obtained from the weighted mean formula. The region of cropping starts from the centroid and is expanded by comparing the surrounding values of centroid with threshold value. The cropped region covers 80% of the total saliency values. Our experiment shows the accomplishment of qualitative image cropping and even the efficiency in terms of time complexity. Moreover, we can demonstrate improvements of our method over recent cropping algorithms on a broad range of images.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

1. INTRODUCTION

Image cropping is very useful in photography, printing industries and so on in many other fields. Its major advantages are improving the composition of image visually, removing the distracted elements from the image, changing the aspect ratio, focusing on a particular part of the image as a subject, changing the orientation as required and zoom in the required part of the image. Whenever there are highly resolution images, cropping can be done and take out the larger amounts of parts in the image that are not considered as the main subject of the image. The importance of the image can be found out using Saliency models. There are various salience detection methods that are useful for object recognition. Some saliency models use weighted dissimilarity measures.

Image cropping can be done in two ways. One is attention based approach and the other is aesthetic based approach. Attention based approach gives importance in finding out the important subject in the image by calculating the attention scores (e.g. saliency [1]) computed over the image. Many of these methods crop the image taking the maximum attention score as reference. Some of these methods use entropy size and distance from the image center as reference [2]. The other approach is aesthetic based that focuses on the attractiveness of the image. It is related to the image composition. These approaches involve in photo quality assessments [3] [4] [5] [6]. More about these approaches can be learned from these papers.

1.1. Motivation

All photo editing tools have an option of cropping which should be done manually by seeing the object in the image which is important. Uses of the image cropping bring motivation to work on it. Some of the advantages are:

1.1.1. Improving the Image composition

To enhance and to draw the attention to main parts of the image some composition guidelines to be followed. If these are followed perfectly, the images don't look like simple snapshots. Some of these guidelines include

a) Rule of thirds
b) View point
c) Depth of the image
d) Unobtrusive background
e) Leading lines
f) Symmetry and Patterns

By having these good guidelines in an image, one can say that the image is having good image composition.

1.1.2. Distracting elements removal

Sometimes images contain shadows and lesser priority objects. These are all not useful and just occupy the space. For removing such parts, image cropping plays a vital role.

1.1.3. Aspect Ratio Change

If there is a requirement of having the same size of images irrespective of taking the image from different types of cameras like DSLR, Micro Four Thirds, 120mm Film cameras and so on, then image cropping method by fixing a standard ratio is required.

1.1.4. Orientation change

Suppose an image is taken in a view that is to be changed to some other view, then cropping allows changing the perspective views. This case works out if an image cannot be taken at a particular place for a particular view.

1.1.5. Zoom in the required part of the image

If a particular part of the image should be find out, then image cropping plays a role in such a way by enlarging the focus of the image. This is useful in crime investigations to find out a particular person in a group of person’s image.

1.1.6. Focusing on the subject

This is the major advantage of image cropping which involves finding out the image importance using some approaches like attention and aesthetic based and the cropped according to the values given by these approaches.

1.2. Proposed Approach

In the current approaches for attention based and automatic image cropping, object is detected using saliency models. Therefore, object detection using Spatially Weighted Dissimilarity Saliency Detection Model is done before cropping.

In the proposed model, we first use the Spatially Weighted Dissimilarity Saliency Detection Model [7] to find out the saliency scores of the image. Using these scores, centroid of the image is found out using the weighted mean formula.

Challenging point in the proposed method is finding out the starting point of cropping. This is done using weighted mean formula. Threshold value is taken at the 21% of the sorted saliency values. Then the region is expanded based on the threshold value. Cropping is done starting from the centroid value and then expanding the region by comparing the surrounding values of centroid with threshold value, Cropping is done till it covers 80 percent of the saliency values. Other challenging point in the proposed model is to do find out the crop quality measures using color distance, color moments, MS Energy and amplitude [8] as parameters.

1.3. Outline of Thesis

The structure of this thesis is as follows: Chapter 2 provides a literature review of different automatic image cropping algorithms using feature extraction and quality classification. Chapter 3 presents our technical approach and discuss the efficiency of the proposed algorithm. Chapter 4 presents the existing approach. Chapter 5 presents the comparison factors used for cropping. Chapter 6 presents the results of comparison between proposed approach and existing approach. Finally, in chapter 7 we draw our conclusions.

CHAPTER II

2. Literature Review

Image cropping is mainly for cropping the important part of the image and removing the distracting elements. This leads to decreasing the storage space. Many worked on this automated image cropping using different algorithms. Some of the most recent works are presented here.

Jianzhou Yan [9] had done the learning the change for Automatic Image Cropping by extracting the features. Feature extraction is done based on the color distance, texture distance, isolation from background, foreground node, shape complexity and sharpness. In our algorithm feature extraction is done for measuring the crop quality. Usually feature extraction is used for content based image retrieval. But Jianzhou Yan used it for measuring crop out and cut through values. Cut through is the chance that the crop boundary passes through a region whereas crop out is the likelihood a region is cropped out of an image. Novelty in this algorithm is usage of Support Vector Machine (SVM) regression technique for computational score assessment which is used for final crop selection. This algorithm may cause foreground detection problem due to which results may be inappropriate in cases which involve obtrusive background.

Masashi Nishiyama [10] had done sensation based photo cropping by using quality classification. This way of cropping involved detecting multiple subject regions using Saliency Map (low-level) method. In our algorithm spatially weighted dissimilarity model is used for getting patches from the image which are used for separating foreground and background of an image. In sensation based photo cropping, regions are acquired using K-means clustering method. Here also, edge, color and blur features are extracted referring to rules of thumb and estimation of posterior probabilities is done. These probabilities are used for matching feature extraction quality. SVM with kernels is used for combining probabilities.

Nehal Jaiswal [11] used Saliency map and SVM for automatic image cropping. A saliency map of Itti is designed for the input image. It identifies salient regions using color, orientation and intensity features. As described earlier our algorithm uses Spatially weighted dissimilarity model for saliency detection. Nehal Jaiswal had done Otsu Thresholding algorithm for separation of wanted and unwanted regions. color and texture features are used for training SVM classifier. Lastly, cropping window is provided using horizontal and vertical profile on the required regions.

Debang Li[14] proposed an algorithm for automatic image cropping using aesthetics aware reinforcement learning. The novelty in this algorithm is reduction of multiple sliding windows which is usually a problem in all existing automatic image cropping algorithms. This is the cause for increase in time for computing cropped image from the original image. To reduce time for cropping window selection, an agent is created which will work according to the pre-defined actions given from the action space. It also follows sequential decision making process but it considers current observation and historical experience. In our algorithm, cropping window is selected based on the saliency values. Initially centroid is found out using weighted mean formula on saliency values and then expansion is done from that point till the window covers 80% of saliency values. Instead of selecting multiple windows and computing quality score for each window using different kind of parameters, our algorithm collects 80% of saliency values which are greater than threshold value. A good photographer takes an image with 80% of subject and 20% of background to get details about the location. Based on this, our algorithm selects the cropping time. Debang Li collected local and global features using View Finding Network. Actions work on adjusting the size, position and shape of the cropping window. Cropping windows are taken into consideration only if the current observation is having higher aesthetic quality score than the existing maximum historical observation. Score is obtained from the View Finding Network.

CHAPTER III

3. Technical Approach & Architecture

In this chapter, we first explain the components of the architecture and then the architecture of the algorithm is shown as a flow diagram.

3.1. Technical Approach

Technical approach mainly composed of three steps. One is finding out the salient region and the second step is to find out the cropping window.

3.1.1. Finding out salient region

This step also results an image exhibiting saliency map. Saliency is a kind of segmentation of image. As segmentation goal says it is a kind of changing the representation of the image for analyzing easily. It is also used for locating objects and to find out the boundaries in an image. It is like labeling the pixels of same characteristics with same label. Some examples of characteristics include texture, shape and color. So, finding out the salient region from the image plays an important role in image cropping which involves removing distracting objects.

The image resulting from this step looks like a gray converted image having high grey level will be shown obviously in saliency map. The model used in our approach is spatially weighted dissimilarity model having three subsequent steps. They are

3.1.1.1 Representing image patches

Assuming an image of size H*W representing H as height of the image and W as the width of the image. Non-overlapping patches are drawn from it. Supposing they are of size k*k pixels. From the above assumptions, the number of patches can be taken as $\lfloor H/k \rfloor * \lfloor W/k \rfloor$. The count of these patches are denoted by L. Each patch is represented as p_i where i=1,2,….,L. As color space has three components, each patch has a length of column vector $3k^2$ represented as f_i. A matrix is formed with all the patches column vectors as A= [f_1 f_2 ……. f_L]

3.1.1.2 Dimensionality Reduction

As A is a feature vector, for dimension reduction Principal Component Analysis is used. For better analyzing the image, this step is done. PCA is used for down sampling. Sampling involves different methods like picking some of the pixels and throwing others, averaging the pixels rather than leaving pixels and some Interpolation methods. In this approach PCA is used for more visualizing significant parts of the image.

f_1,f_2, ……f_L are all subtracted with the average of all the columns

→$f_1=f_1-((f_1+f_2+f_3+ \ldots\ldots f_L)/L)$

Co-similarity matrix is determined as follows

$G=(A^TA) / L^2$

→G is a L*L matrix. Since (L*1) *(1*L) =L*L

Eigen vectors U=[X_1,X_2,X_3,…..X_L]

Where X_i is an eigen vector

U size is d*L

'd' is the biggest eigen value selected

Using the values of eigen vectors, positions are determined for mapping the patches to reduced dimensional space.

3.1.1.3 Spatially Weighted Dissimilarity Evaluation

Dissimilarity evaluation is based on two factors. They are dissimilarities and spatial distance between patches in the reduced dimensional space. As the distance between patches increases, the similarity between them increases. They are inversely proportional to each other. The distance of each patch from the center is considered for saliency evaluation for central bias stated in [3,27]. If the distance from the center to patch is increasing, the decrement of saliency for the patch is done. Using these values, saliency of patch p_i is calculated as follows

$$Saliency(i)=w_2(i)*\sum\{w_1(i,j) * Dissimilarity(i,j)\}$$

Where w_1 (i,j) is defined as

$w_1 (i,j)= 1/(1+dist(p_i,p_j))$

dist((p_i,p_j) is the distance between the center of patches p_i and p_j

Dissimilarity (i,j) $=\sum |x_{si} – x_{sj}|$

$W_2(i)= 1- distToCenter(p_i)/D$

D= $max_j\{distToCenter(p_j)\}$ →normalization factor

DistToCenter is the distance between centers of patch p_i and the original image

3.1.2. Finding out cropping window

To find out the cropping window, there are two important steps. First of all, a threshold value should be set up for the comparison of saliency values. The value is picked out in in such a way the cropping window should cover 80% of saliency values from the original image. To get that threshold value, the saliency values from the Spatially Weighted Dissimilarity Model are sorted and the value located at 20^{th} percentile of all the values is taken as threshold.

Second step is to find out the location from where the 80% of saliency values to be covered. For finding out the location to start expanding, weighted mean formula is used. The location which we get is taken as centroid and the surrounding values are compared in such a way, that the value greater than threshold value is to be in cropped region. The centroid is calculated as follows.

calculate the weighted X value and weighted Y value

Weighted X = $\sum$ i * scaledPixel

Weighted Y = $\sum$ j * scaledPixel

Where

scaledPixel=(Saliency_pixel (i,j) -Min)/(Max-Min)

totalWeight= $\sum$scaledPixel(k)

Max= maximum saliency value of total pixels in the image

Min= minimum saliency value of total pixels in the image

Centroid(x,y)=((Weighted X/totalWeight), (Weighted Y/totalWeight))

After getting the centroid location, the neighboring pixels are compared with threshold value. If one of the neighboring pixel is higher than the threshold, the region expands in the direction of that pixel having saliency value greater than threshold value. This expansion continues until the cropping window cover 80% of the saliency values.

3.2. Architecture of the model

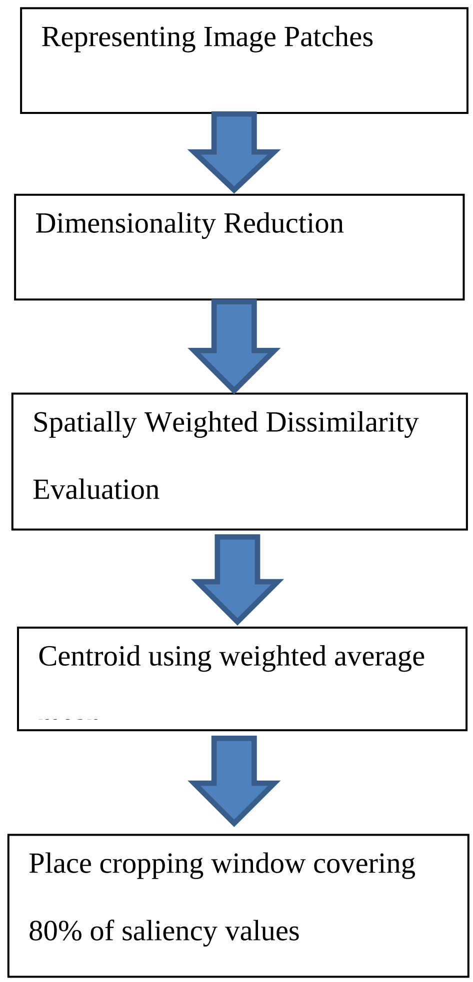


Figure 3.1 Architecture of Proposed Approach

CHAPTER IV

4. EXISTING APPROACH

The comparison of proposed approach is done using an existing algorithm [14]. In this approach, aesthetic aware reinforcement is done using View Finding Network [15]. It is used for finding the good compositions in an image. A set of 14 predefined actions are given to an agent categorized as aspect ratio translation actions, position translation actions, scaling actions and termination action. The cropping windows are adjusted according to the reward obtained using VFN. If the reward is more for the current cropping window than the existing reward for the cropping window in the history then the cropping window is changed to the current location. According to the predefined actions, the selection of cropping windows is done. Suppose agent used action 1 for the selection of a cropping window. If the next cropping window selection is done using action 2. This approach is designed in such a way, the difference between the first cropping window and second cropping window size varies by 0.05 times. These actions are used for adjusting the shape and position of the cropping window by 0.05 times. If the cropping window variation is less than 0.05 times, it is not considered in this approach. The model of this algorithm is as follows.

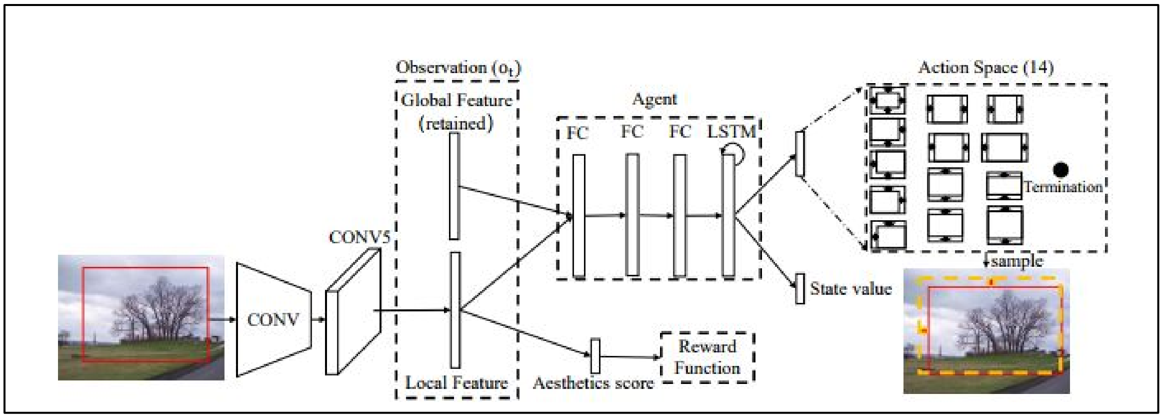


Figure 4.1 Architecture of Existing Approach (Li et al, 2017)

Aesthetics score is the reward which we are talking above. The rewards function is obtained using the following formula.

$$r_{basic_t} = sign(S_{\text{VFN}}(I_t) - S_{\text{VFN}}(I_{t-1}))$$

Where $S_{VFN}(I_t)$ is the score given by the VFN for current cropped image

$S_{VFN}(I_{t-1})$ is the score given by the VFN for last cropped image

If the $r_{basicT}>0$, the current cropped image is considered else last cropped image remains. This step is done for all the sliding windows obtained by agent using predefined actions.

Local and global features are extracted in this approach to find out the quality assessment score. An agent is used for monitoring and performing the cropping. The cropping windows are all placed according to the actions given to the agent. The finalization of the cropped window depends on the reward function.

CHAPTER-V

5. CROP QUALITY MEASUREMENT

The crop quality measurement is done using some parameters. They are HSV histograms, color moments, mean square energy, mean amplitude and area of the image. A brief description about parameters is given below.

5.1. HSV histograms

Color histograms represents the image color distributions. It is used for crop quality measuring metric, because histogram equalization of a color image is used for finding out the image intensity. Separation of color components from intensity is usually done for lighting changes and removing shadows. If an image contains shadow, RGB possesses very different characteristics which will not give accurate values whereas in HSV only value (V) and saturation (S) values may affect but not hue (H) which indicates primary color. In image cropping, removal of shadows plays a major role for which HSV conversion is important. So, HSV histograms values are computed and are compared to the original image using Spearman distance metric. After removing those shadows from the image, the HSV values decreases, due to which the distance value will be more for good crop.

5.2. Color moments

Color moments are used for finding out the similarity measure between images based on the features of color. These can be computed for any color model. They are useful for finding out the shape information of the image. If the color moments score is less, the images are said to be more similar. So, the color moments value should be less for the cropped image. It implies the distance between the original image and cropped image should be more. So, this also can be used as a metric for crop quality.

5.3. Mean square energy

Energy can be used for various purposes in image processing. It can be used for maximizing the performance of object detection and segmentation. It can also be used in image transforms. For characterizing the content of an image, some statistical parameters are used. They are classified as first order, second order and higher order. Among those, second order statistics like mean, dispersion, entropy, means square energy or average energy, kurtosis and skewness are calculated for common features extraction from the image. These are used to get better understanding about images. This value should be less after cropping an image which leads to the distance increase from original image. This is used as crop quality metric to get features comparison between original and cropped image.

5.4. Mean amplitude

An image is defined as a function of a(x,y) with ‘a’ as amplitude of the image. Sometimes image may contain regions of interest as sub images, in that case if an image contains more than one object; the region of interest will be affected by amplitude. Some image processing mechanisms are allowed to do only on the objects in the image rather than on the whole image. In image cropping also, image should be cropped in such a way, the main content of the image should not be disturbed. The mean amplitude value should be less as the cropping should not disturb the content of the image due to which the distance between the original image amplitude and cropped image amplitude should be more. This can also be used as a metric because, after cropping the content should not be disturbed.

5.5. Area of the image

According to the proposed approach, the image should be cropped in such a way it should cover 80% of saliency values. If more part of the image is removed, the area of the image decreases. So, removal of more parts from the image is inversely proportional to the area of the image because removing more parts from the image is nothing but decreasing the size of the image which will decrease pixels count in image. The area of the image is calculated as product of height and width of the image. As more parts are removed, the height and width decreases leading to their product reduction. Sometimes this removal of parts may lead to removing the content from the image which is considered to be bad cropping. The area of the image after cropping should be qualitative having complete important content of the image without loss.

CHAPTER VI

6. RESULTS

A set of expert photographer's images are taken for cropping on the approach we followed. After cropping, the crop quality measures are calculated for each image. Only 100 images were shown for each and every parameter to show the efficient peak point to compare existing versus proposed approaches. Those 100 images were selected using random sampling.

6.1. HSV HISTOGRAM

A graph which is explaining the comparison between the proposed and existing algorithms in terms of differences in histogram values. The values on the Y-axis represents the Spearman distance measure. Calculation of distance between original image HSV histogram value and cropped image HSV histogram value is done for both existing and proposed algorithms. They are drawn in the graph as follows according to the image IDs which are plotted on X-axis.

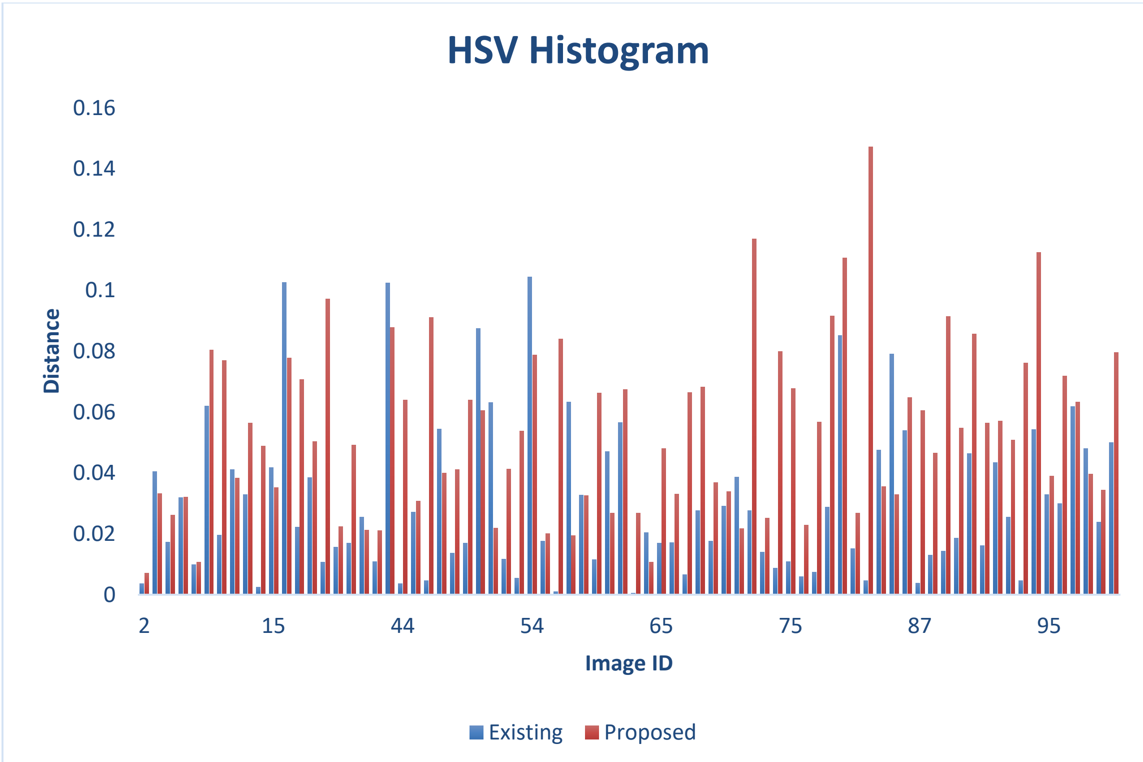
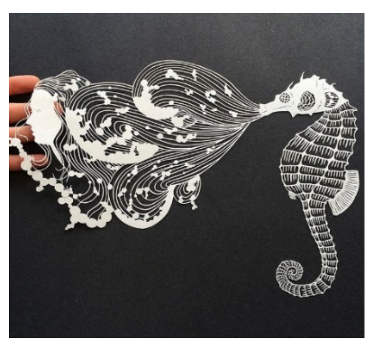


Figure 6.1Graph showing the variation of HSV histograms for existing and proposed approach

According to the HSV histogram parameter graph, the value at image ID 83 is having large difference between the proposed approach and existing approach. The image at image ID 83 is having the following values. The variation can be clearly seen below.



Type: Original Type: Proposed Type: Existing

Size: 720*661

Size: 620*617

Distance: 0.14727

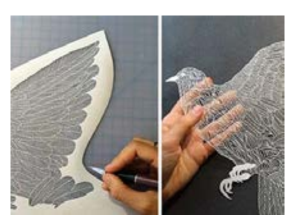
Size: 720*661

Distance: 0.004574

Justification:

From the above images, it can be clearly seen that the size of the image in existing approach is more than the size of the image in the proposed approach. Coming to quality in terms of HSV histogram values, the value of proposed approach should be less than the existing approach because when RGB values are converted to HSV, the removal of shadows and lighting changes of an image are done. After removal of these affects, the HSV values decreases for which the distance between the HSV value of original image and proposed cropped image HSV value increases. When the difference between the original image and existing image is taken in terms of HSV histograms using Spearman distance measure, the value should be lesser than the difference between original image and the proposed image. It is proven in the case of IMAGE ID 83 where 0.004574<0.14727. As it is lesser, the proposed approach is better for this image compared to existing approach.

According to the HSV histogram parameter graph, the value at image ID 54 is having large difference between the proposed approach and existing approach. The image at image ID 54 is having the following values. The variation can be clearly seen below.



Type: Original	Type: Proposed	Type: Existing
Size: 720*360	Size: 581*359	Size: 396*288
	Distance: 0.078836	Distance: 0.104391

Justification:

From the above images, it can be clearly seen that the size of the image in existing approach is less than the size of the image in the proposed approach. Coming to quality in terms of HSV histogram values, the value of proposed approach should be less than the existing approach because when RGB values are converted to HSV, the removal of shadows and lighting changes of an image are done. After removal of these affects, the HSV values decreases for which the distance between the HSV value of original image and proposed cropped image HSV value increases. When the difference between the original image and existing image is taken in terms of HSV histograms using Spearman distance measure, the value should be lesser than the difference between original image and the proposed image. It has got disproved in the case of IMAGE ID 54 where 0.104391>0.078836. As it is greater, the proposed approach is not better for this image compared to existing approach.

6.2. COLOR MOMENTS

A graph which is explaining the comparison between the proposed and existing algorithms in terms of differences in color moments values. The values on the Y-axis represents the Spearman distance measure. Calculation of distance between original image color moments values and cropped image color moments values is done for both existing and proposed algorithms. They are drawn in the graph as follows according to the image IDs which are plotted on X-axis.

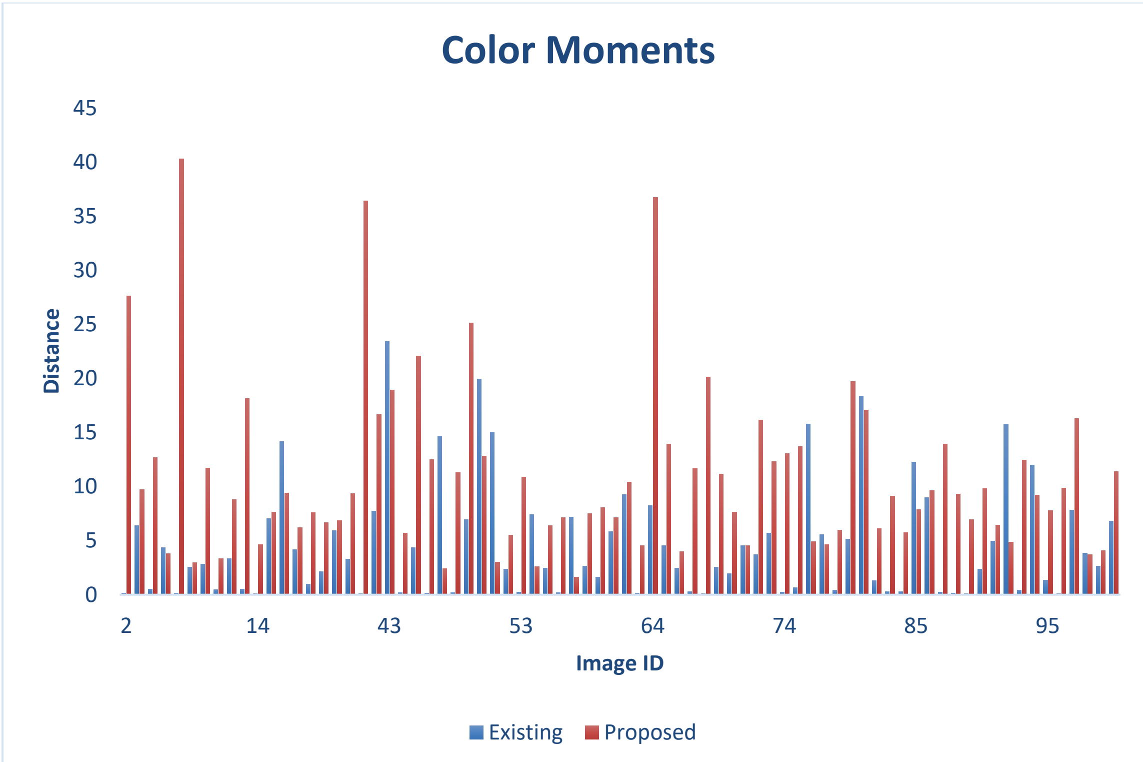


Figure 6.2 Graph showing the variation of color moments for existing and proposed approach

According to the color moments parameter graph, the value at image ID 64 is having large difference between the proposed approach and existing approach. The image at image ID 64 is having the following values. The variation can be clearly seen below.

Type: Original	Type: Proposed	Type: Existing
Size: 1600*1600	Size: 1446*1421	Size: 1520*1600
	Distance: 36.72404	Distance: 8.230808

Justification:

From the above images, it can be clearly seen that the size of the image in existing approach is more than the size of the image in the proposed approach. Coming to quality in terms of color moments values, the shape of the image should not be disturbed. The distraction of the shape information which is obtained using the color moments value should be less when computed for the proposed cropped image. When the difference between the original image and existing image is taken in terms of color moments using Spearman distance measure, the value should be lesser than the difference between original image and the proposed image. It is proven in the case of IMAGE ID 64 where 8.230808<014727. As it is lesser, the proposed approach is better for this image compared to existing approach.

According to the color moments parameter graph, the value at image ID 43 is having large difference between the proposed approach and existing approach. The image at image ID 43 is having the following values. The variation can be clearly seen below.

Type: Original	Type: Proposed	Type: Existing
Size: 1600*1417	Size: 1365*1332	Size: 1360*1276
	Distance: 18.93399	Distance: 23.43045

Justification:

From the above images, it can be clearly seen that the size of the image in existing approach is more than the size of the image in the proposed approach. Coming to quality in terms of color moments values, the shape of the image should not be disturbed. The distraction of the shape information which is obtained using the color moments value should be less when computed for the proposed cropped image. When the difference between the original image and existing image is taken in terms of color moments using Spearman distance measure, the value should be lesser than the difference between original image and the proposed image. It has got disproved in the case of IMAGE ID 43 where 23.43045>18.93399. As it is greater, the proposed approach is not better for this image compared to existing approach.

6.3. MEAN SQUARED ENERGY

A graph which is explaining the comparison between the proposed and existing algorithms in terms of differences in energy values. The values on the Y-axis represents the Spearman distance measure. Calculation of distance between original image mean squared energy values and cropped image mean squared energy values is done for both existing and proposed algorithms. They are drawn in the graph as follows according to the image IDs which are plotted on X-axis.

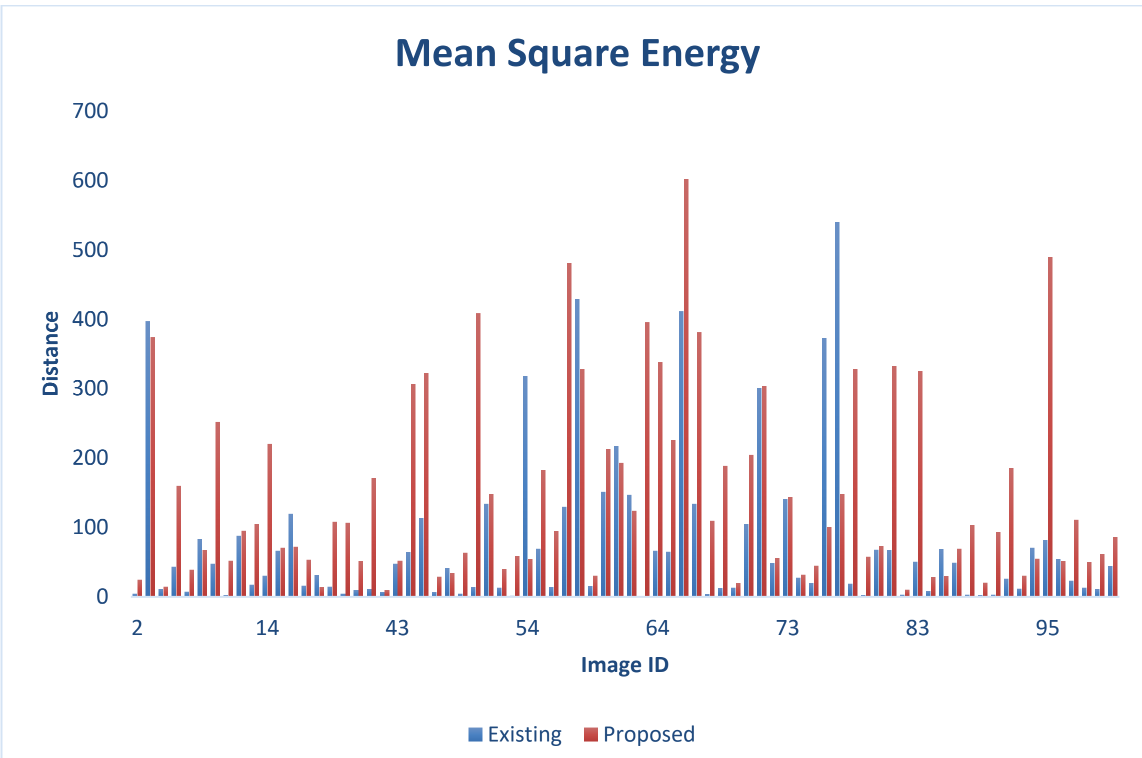


Figure 6.3 Graph showing the variation of Mean squared Energy for existing and proposed approach

According to the mean squared energy parameter graph, the value at image ID 66 is having large difference between the proposed approach and existing approach. The image at image ID 66 is having the following values. The variation can be clearly seen below.

Type: Original

Size: 1024*768

Type: Proposed

Size: 830*761

Distance: 601.5356

Type: Existing

Size: 922*768

Distance: 410.6937

Justification:

From the above images, it can be clearly seen that the size of the image in existing approach is more than the size of the image in the proposed approach. Coming to quality in terms of mean square energy values, better understanding about the image is preferred. The value of mean square energy should be less as it implies that there is better understanding about the image. Understanding in the image include good features are included in the image which are leading to good composition of an image. Those features should not be disturbed after cropping an image. When the difference between the original image and existing image is taken in terms of mean square energy using Spearman distance measure, the value should be lesser than the difference between original image and the proposed image. It is proven in the case of IMAGE ID 66 where 410.6937<601.5356. As it is lesser, the proposed approach is better for this image compared to existing approach.

According to the mean squared energy parameter graph, the value at image ID 77 is having large difference between the proposed approach and existing approach. The image at image ID 77 is having the following values. The variation can be clearly seen below.

Type: Original

Size: 1600*1200

Type: Proposed

Size: 1297*1188

Distance: 147.7859

Type: Existing

Size: 1360*900

Distance: 539.7993

Justification:

From the above images, it can be clearly seen that the size of the image in existing approach is more than the size of the image in the proposed approach. Coming to quality in terms of mean square energy values, better understanding about the image is preferred. The value of mean square energy should be less as it implies that there is better understanding about the image. Understanding in the image include good features are included in the image which are leading to good composition of an image. Those features should not be disturbed after cropping an image. When the difference between the original image and existing image is taken in terms of mean square energy using Spearman distance measure, the value should be lesser than the difference between original image and the proposed image. It has got disproved in the case of IMAGE ID 77 where 539.7993>147.7859. As it is greater, the proposed approach is not better for this image compared to existing approach.

6.4. MEAN AMPLITUDE

A graph which is explaining the comparison between the proposed and existing algorithms in terms of differences in amplitude values. The values on the Y-axis represents the Spearman distance measure. Calculation of distance between original image mean amplitude values and cropped image mean amplitude values is done for both existing and proposed algorithms. They are drawn in the graph as follows according to the image IDs which are plotted on X-axis.

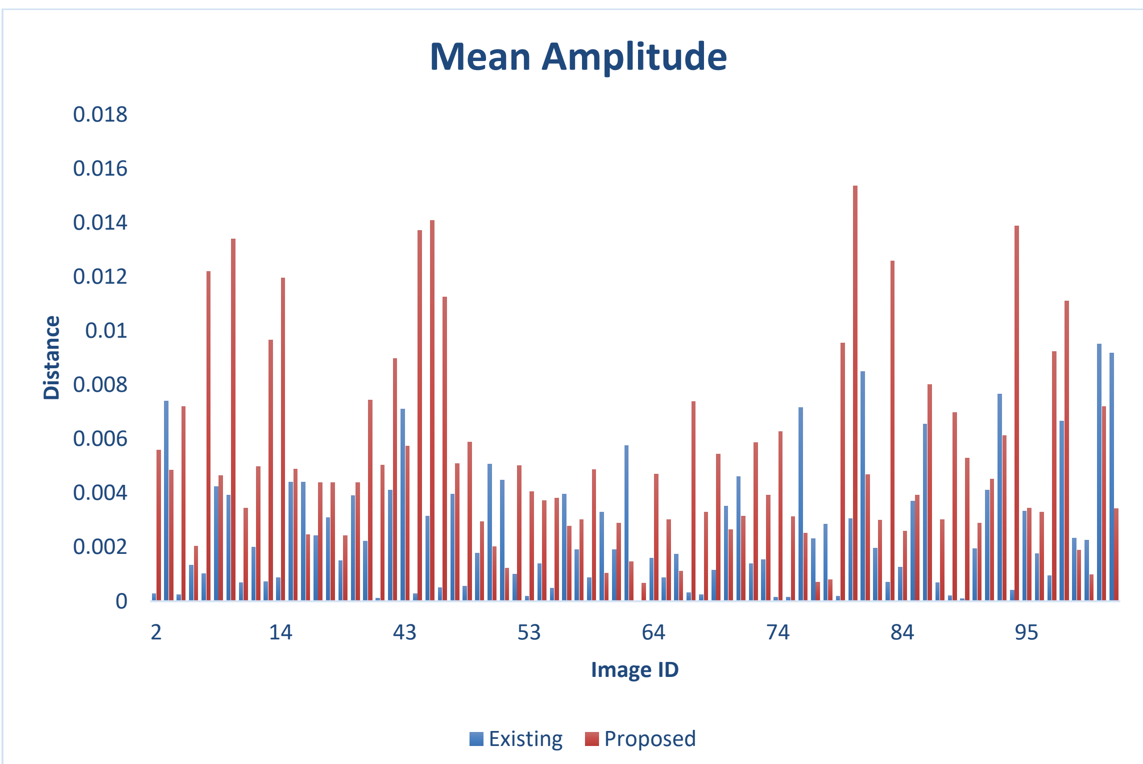
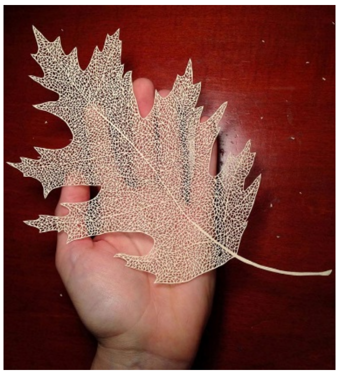
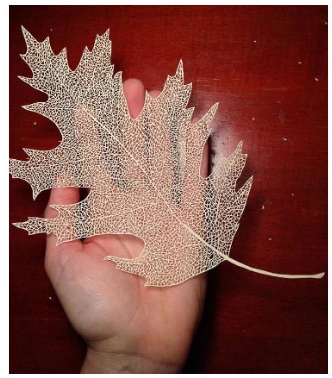


Figure 6.4 Graph showing the variation of Mean amplitude for existing and proposed approach

According to the mean amplitude parameter graph, the value at image ID 80 is having large difference between the proposed approach and existing approach. The image at image ID 80 is having the following values. The variation can be clearly seen below.





Type: Original	Type: Proposed	Type: Existing
Size: 720*790	Size: 671*682	Size: 684*790
	Distance: 0.01537	Distance: 0.003059

Justification:

From the above images, it can be clearly seen that the size of the image in existing approach is more than the size of the image in the proposed approach. Coming to quality in terms of mean amplitude values, the image content should not be disturbed more after cropping. For which the disturbance calculation is done using mean amplitude metric in the image. This value should be less which indicated good cropping. When the difference between the original image and existing image is taken in terms of mean amplitude using Spearman distance measure, the value should be lesser than the difference between original image and the proposed image. It is proven in the case of IMAGE ID 80 where 0.003059<01537. As it is lesser, the proposed approach is better for this image compared to existing approach.

According to the mean amplitude parameter graph, the value at image ID 100 is having large difference between the proposed approach and existing approach. The image at image ID 100 is having the following values. The variation can be clearly seen below.

Type: Original	Type: Proposed	Type: Existing
Size: 1600*1200	Size: 1285*1198	Size: 1520*1020
	Distance: 0.007207	Distance: 0.009513

Justification:

From the above images, it can be clearly seen that the size of the image in existing approach is more than the size of the image in the proposed approach. Coming to quality in terms of mean amplitude values, the image content should not be disturbed more after cropping. For which the disturbance calculation is done using mean amplitude metric in the image. This value should be less which indicated good cropping. When the difference between the original image and existing image is taken in terms of mean amplitude using Spearman distance measure, the value should be lesser than the difference between original image and the proposed image. It has got disproved in the case of IMAGE ID 100 where 0.009513>0.007207. As it is greater, the proposed approach is not better for this image compared to existing approach.

6.5. FEATURE VECTOR COMPARISON

A graph which is explaining the comparison between the proposed and existing algorithms in terms of differences in feature vectors values. Feature vector is formed by combining HSV histogram, color moments, mean squared energy and mean amplitude. As all the individual values should be less for a good cropping, the aggregation of these values should be obviously less when compared to the existing approach. It implies that the distance of feature vectors from original to proposed cropped image versus original to existing cropped image should be more. The values on the Y-axis represents the Spearman distance measure. Calculation of distance between original image feature vector value and cropped image feature vector value is done for both existing and proposed algorithms. They are drawn in the graph as follows according to the image IDs which are plotted on X-axis.

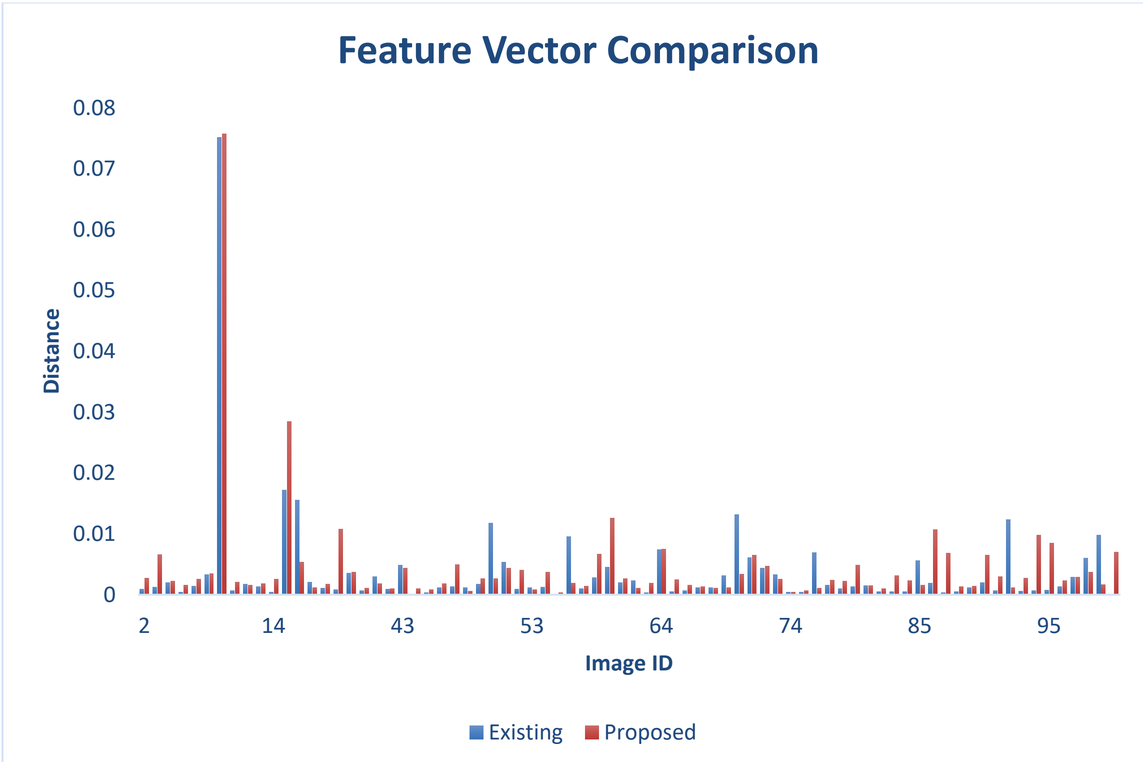
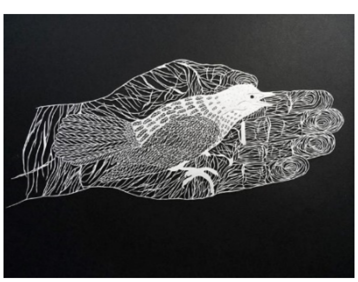
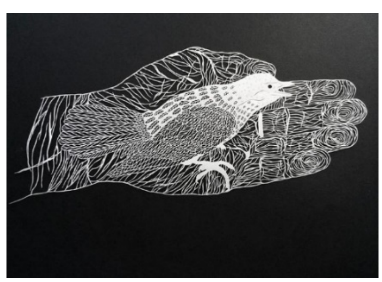


Figure 6.5 Graph showing the variation of Feature Vectors for existing and proposed approach

According to the Feature vector comparison graph, the value at image ID 10 is having large difference between the proposed approach and existing approach. The image at image ID 10 is having the following values. The variation can be clearly seen below.


Type: Original Type: Proposed Type: Existing

Size: 720*540	Size: 720*513	Size: 585* 540
	Distance: 0.075759	Distance: 0.07516

Justification:

From the above images, it can be clearly seen that the size of the image in existing approach is more than the size of the image in the proposed approach. Coming to quality in terms of feature vector values, as all the individual values should be less for a good cropping, the aggregation of these values which are formed as feature vectors should be obviously less when compared to the existing approach. It implies that the distance of feature vectors from original to proposed cropped image versus original to existing cropped image should be more. When the difference between the original image and existing image is taken in terms of feature vector using Spearman distance measure, the value should be lesser than the difference between original image and the proposed image. It is proven in the case of IMAGE ID 10 where 0.07516<0.075759. As it is lesser, the proposed approach is better for this image compared to existing approach.

6.6. PERCENTAGE OF SALIENCY

A graph which is explaining the comparison between the proposed and existing algorithms in terms of percentages of saliency values. Saliency measure is used to find out the region of interest in the image. If the more part of important content is removed from the image, the saliency measure decreases. Those are inversely proportional to each other. The values on the Y-axis represents the percentage of saliency measure. They are drawn in the graph as follows according to the image IDs which are plotted on X-axis. According to good photography rules, images should contain maximum of portion as important content in the image without considering background distracting objects. According that rule, about 500 images are taken and are cropped using existing and proposed algorithms. After that the calculation of saliency present in each and every cropped image is calculated and plotted graph according to those values.

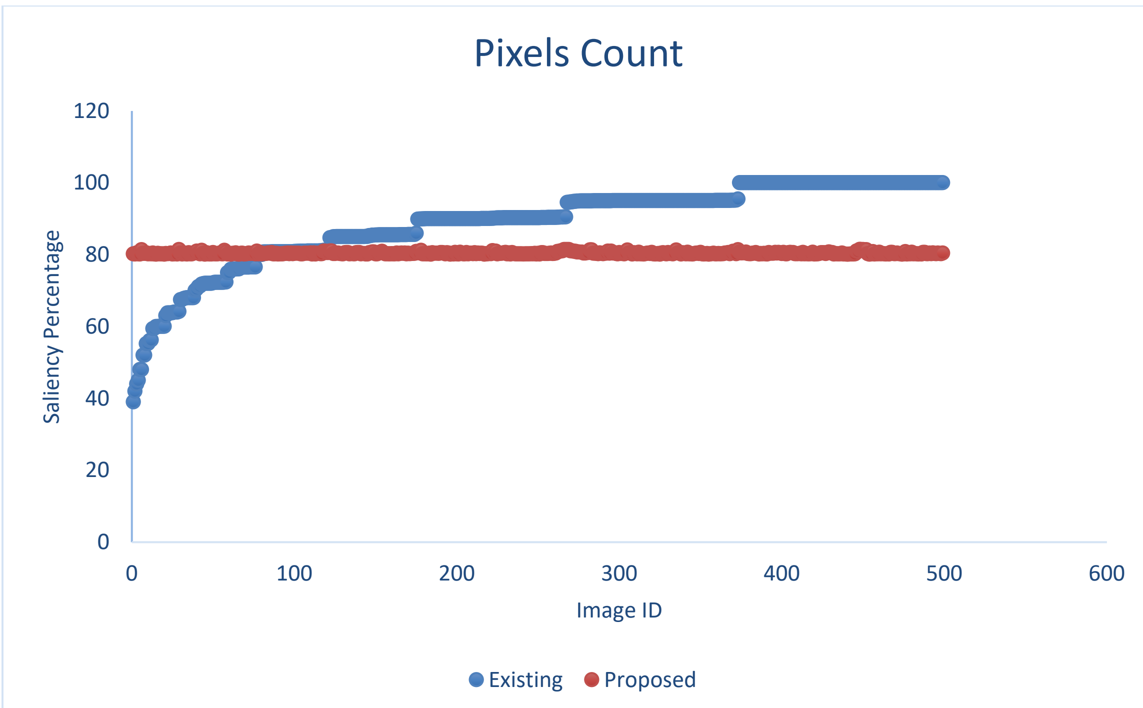
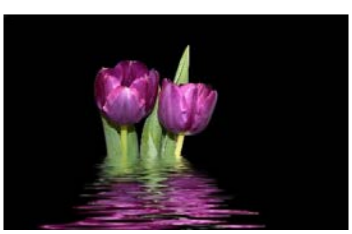
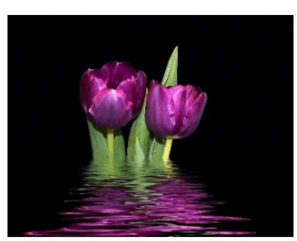
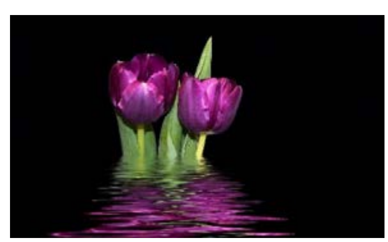


Figure 6.6 Graph showing the variation of percentage of Saliency for existing and proposed approach

According to the Percentage of Saliency graph, the value at image ID 69 is having no difference between the proposed approach and existing approach. The image at image ID 69 is having the following values. The variation can be clearly seen below.







Type: Original	Type: Proposed	Type: Existing
Size: 1280*800	Size: 1029*799	Size: 1280*760
Saliency%: 100	Saliency %: 80.568	Saliency%: 80.524

According to the Percentage of Saliency graph, the value at image ID 391 is having more difference between the proposed approach and existing approach. The image at image ID 391 is having the following values. The variation can be clearly seen below.

Type: Original	Type: Proposed	Type: Existing
Size: 600*399	Size: 485*398	Size: 600*399
Saliency%: 100	Saliency%: 81.645	Saliency%: 100

6.7. OVERALL COMPARISON

The comparison of all the parameters for all the images are displayed in the below graph using histograms. The graph clearly explains that the number of images out of 500 are having more efficiency. Consider hsv histogram value showing 321 images are having more efficiency in proposed approach than existing approach whereas 134 images are having more efficiency in existing approach than proposed approach. Similarly, all parameters are calculated and drawn graph according to the results in the following table.

Parameter Names	Proposed	Existing
HSV	321	179
Color Moments	353	147
Amplitude	358	142
Energy	329	171
Pixels Count	390	110

Table 6.1 Performance in terms of number of images for all parameters

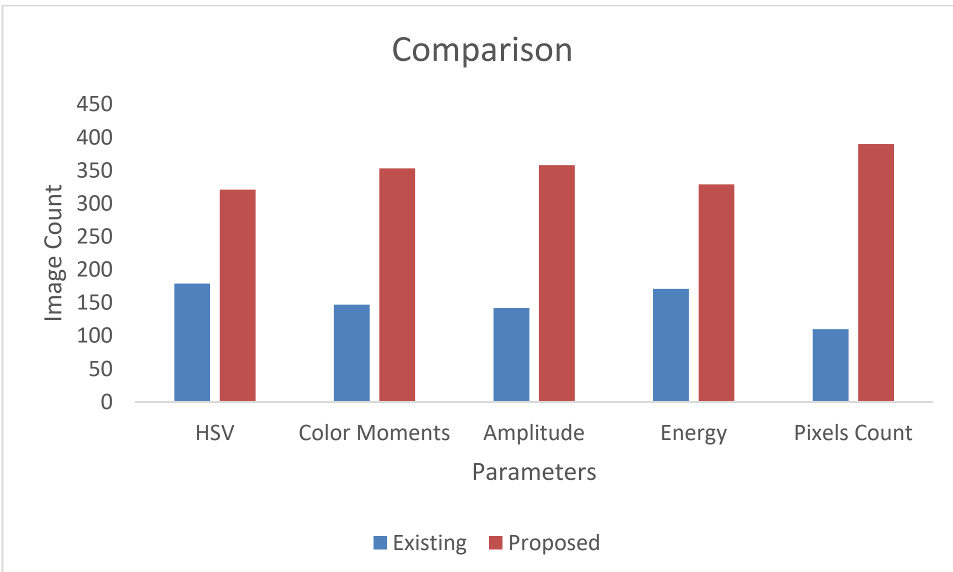


Figure 6.7 Graph showing the comparison of all parameters for existing and proposed approach

CHAPTER VII

7. DISCUSSION

Similar to this weighted dissimilarity model, there are many saliency algorithms [16] [17], which are used for object detection. Therefore, for removal of distracting parts in the image, one should know about the region of interest in the image which is founded out using saliency models. In initial phase, keeping this in mind we thought of finding out a model using saliency detection. Cropping is done by many people using feature extraction [18] which is time taking and requires many mathematical computations. So, we thought of finding an easy way of cropping part from the image. In the process of exploring many papers [19] [20] we got an idea of finding out the part where the cropping should be started.

Knowing about the similarity measures used among images to compare, an idea of using those measures for crop quality arose. For content based image retrieval [21] [22] [23], features are extracted and a similarity metric is found out using distances like city block, minkowski, spearman, Euclidean and so on. We used spearman distance for finding out difference between original image and cropped image obtained for proposed approach. Formula behind this distance is as follows

Spearman Rank Correlation measures the correlation between two sequences of values. The two sequences are ranked separately and the differences in rank are calculated at each position, *i*. The distance between sequences ***X*** = (***X****1*, ***X****2*, etc.) and ***Y*** = (***Y****1*, ***Y****2*, etc.) is computed using the following formula:

$$1 - \frac{6 \sum_{i=1}^{n} (rank(X_i) - rank(Y_i))^2}{n(n^2 - 1)}$$

Where ***Xi*** and ***Yi*** are the i^{th} values of sequences ***X*** and ***Y*** respectively.

The range of Spearman Correlation is from -1 to 1. Spearman Correlation can detect certain linear and non-linear correlations.

CHAPTER VIII

8. CONCLUSION

A model which can easily crop the images without disturbing the content is to be find out without using feature extraction is the main motto of this work. We proposed a new way of automatic image cropping. Particularly, in our method saliency detection is used as many methods but cropping window is formed using a point obtained by weighted mean formula. The location which we obtain is using saliency values. So, it is the centroid of the region of interest in the image rather than the centroid of the image which is a good point to start expanding the region to be cropped. From then, our approach crops the image till it covers 80% of saliency values.

The most important steps in our approach are: First, finding out the most important region in the image. This is done using weighted spatially dissimilarity model. A perfect saliency model should be selected as it is the one and only one measure used for finding out the location of the saliency where the cropping should be started. Second, we found out the centroid using weighted mean formula on saliency values obtained in the first step. Third step is to expand the region of cropping from the point we got in the foist step.

In the future, we hope to perform on different dataset containing bad photographs and do qualitative analysis on those also. We also hope to perform comparing the crop quality using more distance metrics like correlation, city block, L1 distance, cosine correlation, minkowski and chevbyshev. We also hope to perform comparison of these distance metrics on the parameters used like hsv histogram, color moments, mean squared energy and mean amplitude.

REFERENCES

[1] L. Itti and C. Koch. A model of saliency based visual attention of rapid scene analysis. IEEE Trans. PAMI, 20(11):1254–1259, 1998.

[2] M. Ma and J. K. Guo. Automatic image cropping for mobile devices with built-in camera. In Consumer Communication and Networking, pages 710–711, 2004.

[3] W. Luo, X. Wang, and X. Tang. Content-based photo quality assessment. In ICCV, pages 2206 –2213, 2011.

[4] Y. Luo and X. Tang. Photo and video quality evaluation: Focusing on the subject. In ECCV, pages III:386–399, 2008.

[5] Y. Ke, X. Tang, and F. Jing. The design of high-level features for photo quality assessment. In CVPR, 2006.

[6] X. Tang, W. Luo, and X. Wang. Content-based photo quality assessment. IEEE Transactions on Multimedia, 2013.

[7] L. Duan, C. Wu, J. Miao, L. Qing, and Y. Fu, “Visual saliency detection by spatially weighted dissimilarity,” in Proc. IEEE Conf. CVPR, Jun. 2011, pp. 473–480

[8] R.S. Choras, Image feature extraction techniques and their applications for CBIR and biometrics systems, Int. J. Biol. Biomed. Engrg. 1 (1) (2007).

[9] J. Yan, S. Lin, S. B. Kang, and X. Tang. Learning the change for automatic image cropping. In IEEE CVPR, 2013.

[10] M. Nishiyama, T. Okabe, Y. Sato, and I. Sato. Sensation-based photo cropping. In ACM MM, 2009

[11] N. Jaiswal and Y. K. Meghrajani, "Saliency based automatic image cropping using support vector machine classifier," in Proc. Int. Conf. Innovations Information, Embedded and Communication Systems, 2015, pp. 1–5.

[12] Subhabrata Bhattacharya, Rahul Sukthankar, and Mubarak Shah. A framework for photo-quality assessment and enhancement based on visual aesthetics. In ACM Multimedia, pages 271–280, 2010.

[13] B.S. Manjunath and W.Y. Ma, ªTexture Features for Browsing and Retrieval of Image Data,º IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18, no. 8, pp. 837-842, Aug. 1996

[14] D. Li, Huikai Wu and Kaiqi Huang A2-RL: Aesthetics Aware Reinforcement learning for Aytomatic Image Cropping in arXiV, Sep. 2017

[15] Chen, Y.-L.; Klopp, J.; Sun, M.; Chien, S.-Y.; and Ma, K.-L. 2017b. Learning to compose with professional photographs on the web. arXiv preprint arXiv:1702.00503.

[16] E. Ardizzone, A. Bruno, and G. Mazzola, “Saliency based image cropping,” in Proc. Int. Conf. Image Analysis and Processing, 2013, pp. 773–782.

[17] J. Zhang and S. Sclaroff, “Saliency detection: A Boolean map approach,” in Proc. IEEE ICCV, Dec. 2013, pp. 153–160.

[18] Gaurav Kumar, and Pradeep Kumar Bhatia,“A Detailed Review of Feature Extraction in Image Processing System” International conference an advanced computing and communication Technologies, 2014.

[19] F. Stentiford, “Attention based auto image cropping,” in Proc. 5th Int. Conf. Computer Vision Systems, Bielefeld, Germany, 2007.

[20] V. Gopalakrishnan, Y. Hu and D. Rajan. Random walks on graphs to model saliency in images. CVPR, 2009.

[21] V. N. Gudivada and J. V. Raghavan, IEEE Computer (Special Issue on Content-Based Image Retrieval Systems), vol. 28, Sept. 1995.

[22] H., MICHOUX, N., BANDON, D., AND GEISSBUHLER, A. 2004. A review of content-based image retrieval systems in medical applications—Clinical benefits and future directions. Int. J. Medical Inf. 73, 1, 1– 23.

[23] A. W., WORRING, M., SANTINI, S., GUPTA, A.,, AND JAIN, R. 2000. Content-based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Mach. Intell. 22, 12, 1349–1380.

VITA

Mounika Ramineni

Candidate for the Degree of

Master of Science

Thesis: SALIENCY BASED AUTOMATED IMAGE CROPPING

Major Field: Computer Science

Education:

Completed the requirements for the Master of Science/Arts in Computer Science major at Oklahoma State University, Stillwater, Oklahoma in December, 2017.

Completed the requirements for the Bachelor of Science/Arts in Information Technology major at MVGR College of Engineering, Vizianagaram, India in 2015.

Experience:

Worked as Graduate Assistant Web Designer at ITLE, OSU